entropy

MDPI

Article

Fractional Jensen–Shannon Analysis of the Scientific Output of Researchers in Fractional Calculus

José A. Tenreiro Machado [1,†] and António Mendes Lopes [2,*,†]

1 Department of Electrical Engineering, Institute of Engineering, Polytechnic of Porto, Rua Dr. António Bernardino de Almeida 431, 4249-015 Porto, Portugal; jtm@isep.ipp.pt
2 UISPA–LAETA/INEGI, Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal
* Correspondence: aml@fe.up.pt; Tel.: +351-913-499-471
† These authors contributed equally to this work.

Academic Editor: Kevin H. Knuth
Received: 9 March 2017; Accepted: 15 March 2017; Published: 17 March 2017

Abstract: This paper analyses the citation profiles of researchers in fractional calculus. Different metrics are used to quantify the dissimilarities between the data, namely the Canberra distance, and the classical and the generalized (fractional) Jensen–Shannon divergence. The information is then visualized by means of multidimensional scaling and hierarchical clustering. The mathematical tools and metrics allow for direct comparison and visualization of researchers based on their relative positioning and on patterns displayed in two- or three-dimensional maps.

Keywords: science metrics; fractional calculus; Canberra distance; generalized entropy; Jensen–Shannon divergence; multidimensional scaling; hierarchical clustering

1. Introduction

Measuring the scientific output of researchers, namely, productivity (quantity) and/or visibility (impact of citation), is important in many circles, such as universities, journals, funding agencies, promotion committees and employers [1].

The h-index was proposed in 2005 by the physicist Jorge E. Hirsch to measure the scientific output of individual researchers [2,3]. A researcher has index $h \in \mathbb{N}$ if h is the largest number such that his h most cited publications have at least h citations each. For determining h, we can adopt a simple procedure. Firstly, we sort the number of citations per publication by decreasing order for obtaining the citation profile $C = \phi(k)$, where $k \in \mathbb{N}$ represents the rank, and, afterwards, for the array $H = \min\left(\phi\left(k\right), k\right)$, we have $h = \max(H)$. In practice, h is close to the intersection of $C = \phi(k)$ and $C = k$. The h-index has a time memory since it captures the accumulation of citations.

In a short period of a few years, Hirsch's index has been accepted in many fields as a criterion for establishing rankings [1]. The h-index has the advantage of incorporating both the quantity and visibility of publications in a single-number criterion [4,5]. Moreover, the index is equally robust to rarely and frequently cited works [6–8]. The domain of application of the h-index surpassed its original purpose [1,6,9] and was adopted for measuring collective scientific output [5,10,11], evaluating the scientific impact of journals [12,13], and quantifying how much work was done in a given topic or compound [14].

Some authors pointed out several disadvantages of the h-index, noting that, like any other one-parameter index, it withdraws the multidimensional nature of scientific output [5]. Others identified additional shortcomings [15], namely its inability to differentiate between active and inactive researchers [16], its sensitivity to long scientific careers [17] and to discipline-dependent citation profiles [18,19], or its difficulty to reflect the role of co-authorship [20,21]. Those limitations led to the proposal of complementary, or alternative, indices to measure scientific output [2,22–24]. Some variations are the g and h^2 indices [23,25], which give more weight to highly cited publications; the e-index that tries to differentiate between researchers with similar h, but different citation profiles [26]; the $hI, norm$-index [27] that seeks to include the effects of co-authorship, first dividing the number of citations by the number of authors of each publication and then calculating the h-index of the normalized citation counts; the R and AR indices [28], where the first measures the h-core's citation intensity, while the second takes the age of publications into account; and the h_{rat}-index [29] that introduces more granularity on the measure than the original h-index, among others [26,30–32].

Fractional Calculus (FC) generalizes the classical differential operations to non-integer orders [33–35]. The area of FC dates back to the year 1695, with the celebrated correspondence between l'Hôpital and Leibniz about the meaning, and apparent paradox, of an n-order time derivative of a function, $f(t)$, $\frac{d^n f(t)}{dt^n}$, for $n = \frac{1}{2}$. However, it was only in the last decades that FC was recognized as playing an important role in modeling and control of many important physical phenomena, and emerged as a key tool in the area of dynamical systems. Nowadays, the FC community is composed of many researchers in different scientific fields, namely, mathematics, physics, biology, finance and geophysics [36–41].

In this paper, different metrics are used for processing citation profiles, namely, the Canberra distance, and the classical and fractional (generalized) Jensen–Shannon divergence. The information is visualized using multidimensional scaling (MDS) and hierarchical clustering (HC) for comparing the scientific output of FC researchers. The MDS and HC generate maps of points in two- and three-dimensional space that represent researchers according to their scientific production. The relative positioning of the points and the emerging patterns allow for a direct interpretations of the results.

In this line of thought, the paper is organized as follows. Sections 2 and 3 present the dataset and the mathematical background, respectively. Section 4 processes the data and discusses the results. Finally, Section 5 draws the main conclusions.

2. The Dataset

We consider 100 researchers in the area of FC from 35 countries. Their geographic origin is summarized in Table 1. We tackle data from the Thomson Web of Science database (http://apps.webofknowledge.com/), retrieved on 25 January 2017.

Researchers with identical names and researchers that use different short names in their publications may pose difficulties in the searching process. For minimizing errors caused by counting incorrectly the number of publications and/or citations, we adopt a combination of several searching fields, namely, the author name, address, and affiliation. In the experiments shown in the following sections, we identify researchers by a two-letter code.

Table 1. Geographic origin and number of Fractional Calculus (FC) researchers considered in this study.

Country	Number	Country	Number	Country	Number
Algeria	1	Greece	1	Russia	6
Australia	1	Hungary	1	Serbia	2
Austria	1	India	5	Singapore	1
Belgium	1	Iran	2	Slovak Republic	2
Brazil	1	Italy	8	South Africa	1
Bulgaria	2	Japan	1	Spain	7
Canada	3	Jordan	1	Switzerland	1
Chile	1	Mexico	1	Turkey	3
China	4	Netherlands	1	UA Emirates	3
Egypt	2	Poland	4	UK	1
France	6	Portugal	6	USA	13
Germany	5	Romania	1		

The data available allows for determining different indices quantifying scientific output. Figure 1 depicts the charts that illustrate the h, g, h^2 and $hI, norm$ indices for one researcher.

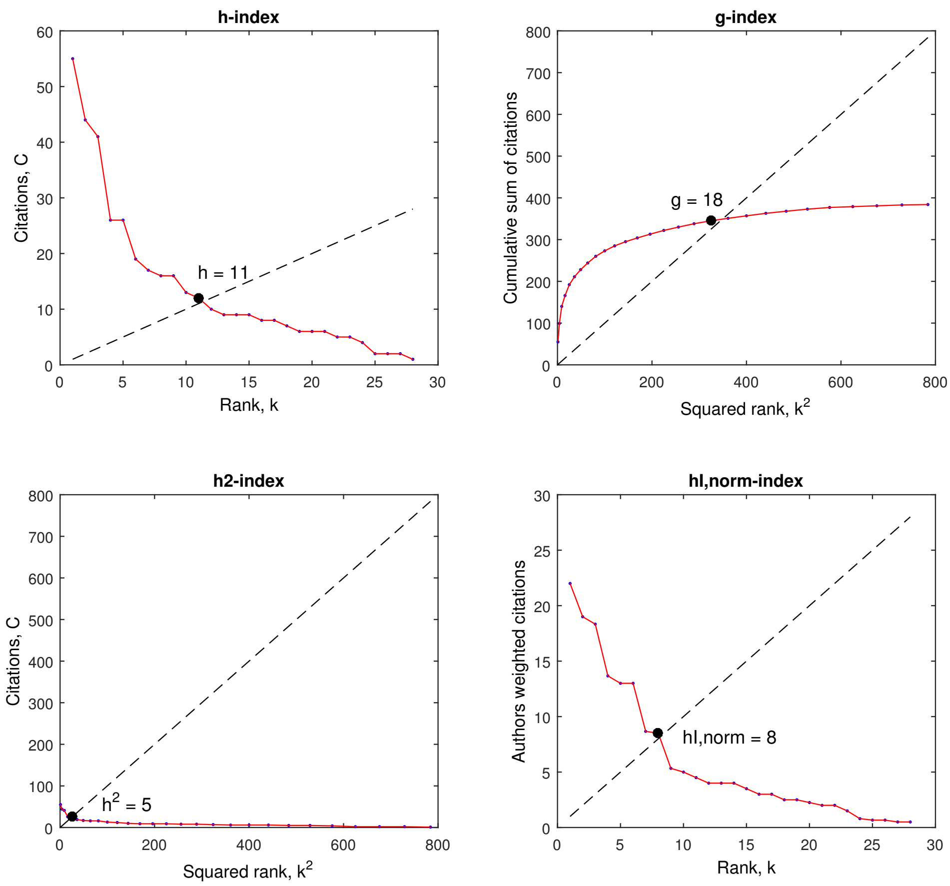

Figure 1. Graphs illustrating the h, g, h^2 and $hI, norm$ indices for one researcher.

In general, the various indices are correlated with each other to some extent. Figure 2 shows the relationships between the indices h, g, h^2 and $hI, norm$ for a group of 100 researchers in FC.

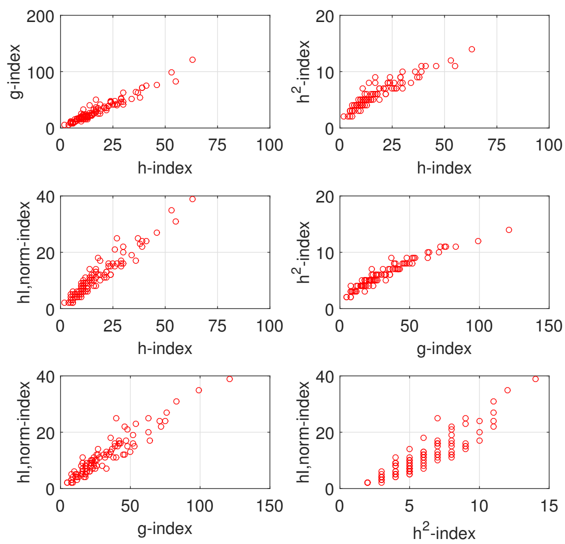

Figure 2. Graphs illustrating the relationships between the indices h, g, h^2 and $hI, norm$ for a group of 100 researchers in FC.

3. Mathematical Background

This section introduces the mathematical background necessary for processing the data, namely, the Canberra distance, the classical and fractional Jensen–Shannon divergence, and the MDS and HC techniques.

3.1. The Canberra Distance

The Canberra distance was proposed, and latter modified, by Lance and Williams [42,43]. Given two points in a K-dimensional space, $X = (x_1, \cdots, x_K)$ and $Y = (y_1, \cdots, y_K)$, the Canberra distance between X and Y is given by:

$$CD(X,Y) = \sum_{k=1}^{K} \frac{|x_k - y_k|}{|x_k| + |y_k|}. \tag{1}$$

Equation (1) is a metric often used for quantifying data scattered around an origin. The Canberra distance has several interesting properties, namely, it is unitary when the arguments are of the opposite sign, biased for measures around the origin, and highly sensitive for values close to zero.

3.2. The Classical and Fractional Jensen–Shannon Divergence

In information theory, the information content of event k with probability of occurrence $p(k)$ is given by:

$$I[p(k)] = -\ln p(k). \tag{2}$$

Recently, inspired by FC, the concept of information content of order $\alpha \in \mathbb{R}$ was proposed as [44,45]:

$$I_\alpha\left[p(k)\right] = D^\alpha I\left[p(k)\right] = -\frac{p(k)^{-\alpha}}{\Gamma\left(\alpha+1\right)}\left[\ln p(k) + \psi\left(1\right) - \psi\left(1-\alpha\right)\right], \tag{3}$$

where $\Gamma(\cdot)$ and $\psi(\cdot)$ denote the gamma and digamma functions, respectively.

The Jensen–Shannon divergence measures the distance between two probability distributions, P and Q [46], and represents a symmetrical and smoothed version of the Kullback–Leibler divergence (or relative entropy), given by:

$$KLD\left(P \parallel Q\right) = \sum_k p(k) \ln \frac{p(k)}{q(k)}. \tag{4}$$

Therefore, we have:

$$JSD\left(P \parallel Q\right) = \frac{1}{2}\left[KLD\left(P \parallel M\right) + KLD\left(Q \parallel M\right)\right], \tag{5}$$

where $M = \frac{P+Q}{2}$ is a mixture distribution.

Alternatively, we may write:

$$JSD\left(P \parallel Q\right) = \frac{1}{2}\left[\sum_k p(k) \ln p(k) + \sum_k q(k) \ln q(k)\right] - \sum_k m(k) \ln m(k), \tag{6}$$

which, using Equation (3), leads to the fractional (generalized) Jensen–Shannon divergence:

$$JSD_\alpha (P \parallel Q) = \frac{1}{2}\sum_k p(k)\left\{\frac{p(k)^{-\alpha}}{\Gamma(\alpha+1)}\left[\ln p(k)+\psi(1)-\psi(1-\alpha)\right]\right\} + \frac{1}{2}\sum_i q(k)\left\{\frac{q(k)^{-\alpha}}{\Gamma(\alpha+1)}\left[\ln q(k)+\psi(1)-\psi(1-\alpha)\right]\right\} - \sum_k m(k)\left\{\frac{m(k)^{-\alpha}}{\Gamma(\alpha+1)}\left[\ln m(k)+\psi(1)-\psi(1-\alpha)\right]\right\}. \quad (7)$$

For $\alpha = 0$, we obtain JSD, as defined in Equation (6).

3.3. Multidimensional Scaling

MDS is a computational technique for clustering and visualizing data [47]. In a first phase, given s items and a measure of dissimilarity, an $s \times s$ symmetric matrix, $\mathbf{\Delta} = [\delta_{ij}]$, $(i,j) = 1, \cdots, s$, of item-to-item dissimilarities, is calculated. Matrix $\mathbf{\Delta}$ represents the input information for starting the MDS numerical scheme. The MDS rational is to assign points for representing items in a multidimensional space and to try to reproduce the measured dissimilarities, δ_{ij}. In a second phase, MDS evaluates different configurations for maximizing some fitness function, arriving at a set of point coordinates (and, therefore, to a symmetric matrix of distances $\mathbf{D} = [d_{ij}]$ that represent the reproduced dissimilarities) that best approximates δ_{ij}. A common fitness function is the raw stress:

$$\mathcal{S} = \left[d_{ij} - f(\delta_{ij})\right]^2, \quad (8)$$

where $f(\cdot)$ indicates some type of transformation.

The MDS interpretation is based on the patterns of points that can be visualized in the map generated. Therefore, the information retrieval is not based on the point coordinates or the geometrical form of the clusters, and we can rotate or translate the map because the distances remain identical.

The "quality" of the MDS map can be assessed by means of the stress and Shepard plots. The stress plot represents $\mathcal{S}$ versus the number of dimensions m of the MDS map. The plot $\mathcal{S}(m)$ is a monotonic decreasing chart and the chosen value of m is a compromise between low values of $\mathcal{S}$ and m. The Shepard diagram, for a particular value m, compares d_{ij} and δ_{ij}. A narrow scatter around the 45 degree line represents a good fit between d_{ij} and δ_{ij}.

3.4. Hierarchichal Clustering

Clustering is a data analysis technique [48] that groups similar items. In HC, two possible iterative strategies generate a hierarchy of clusters, namely, the (i) agglomerative and the (ii) divisive clustering. With (i), each item starts in its own cluster and the algorithm merges the two most similar clusters until there is one single cluster. With (ii), all of the items start in a single cluster and the algorithm removes the "outsiders" from the least cohesive cluster, until each item is in its own cluster. In both cases, a linkage criterion is required, which is a function of the distances between pairs of items for quantifying the dissimilarity between clusters. For two clusters, R and S, the distance $d(x_R, x_S)$ between items $x_R \in R$ and $x_S \in S$ is based on metrics such as the maximum, minimum and average linkages given by [49]:

$$d_{max}(R,S) = \max_{x_R \in R, x_S \in S} d(x_R, x_S), \quad (9)$$

$$d_{min}(R,S) = \min_{x_R \in R, x_S \in S} d(x_R, x_S), \quad (10)$$

$$d_{ave}(R,S) = \frac{1}{\parallel R \parallel \parallel S \parallel} \sum_{x_R \in R, x_S \in S} d(x_R, x_S). \quad (11)$$

After using one of the algorithms, the results of HC are presented in a graphical object such as a dendrogram or a hierarchical tree.

To assess the "quality" of the clustering, the cophenetic correlation (CC) coefficient is used [50]. The CC gives a measure of how well the generated graphical object preserves the original pairwise distances. If the clustering is successful, the links between items in the graphical object have a strong correlation with those in the original data set. The closer the CC value to 1, the better the clustering result. The quality assessment is plotted in a Shepard diagram that compares the original and the cophenetic distances. As for the MDS, a good clustering leads to a layout of points close to the 45 degree line.

4. Data Analysis and Results

In this section, the Canberra distance and the classical and fractional Jensen–Shannon divergence are adopted for quantifying dissimilarities between citation profiles of $s = 100$ researchers in FC. The dissimilarities are processed by the MDS and HC for generating maps of items, representing researchers according to their scientific output.

4.1. Comparing and Visualizing Scientific Output by Means of MDS

Given the citation profiles (ϕ_i, ϕ_j) of researchers i and j, respectively, we first calculate a 100×100 symmetric matrix $\mathbf{\Delta} = [\delta_{ij}]$, $(i, j) = 1, ..., 100$, where δ_{ij} denotes either $CD(\phi_i, \phi_j)$, given in Equation (1), $JSD(\phi_i, \phi_j)$, defined in Equation (6), or $JSD_\alpha(\phi_i, \phi_j)$, as in Equation (7). For CD, the parameter K represents the length of the larger citation profile of the pair (i, j). The smaller profile has to be filled with trailing zeroes for obtaining equal length ϕ_i and ϕ_j vectors. For JSD and JSD_α, the probabilities are approximated by $p(k) = \frac{\phi(k)}{\sum_k \phi(k)}$. The matrix $\mathbf{\Delta}$ then feeds the MDS algorithm.

Figures 3–5 depict the three-dimensional maps generated by the MDS with CD, JSD and JSD_α ($\alpha = 0.7$). The value of $\alpha = 0.7$ was chosen to obtain good discrimination between items [51]. For all cases, we observe that the points representing researchers form similar patterns, namely, those on the right-hand side of the charts. Nevertheless, any other possible patterns (in case they exist) are hidden by the large number of points.

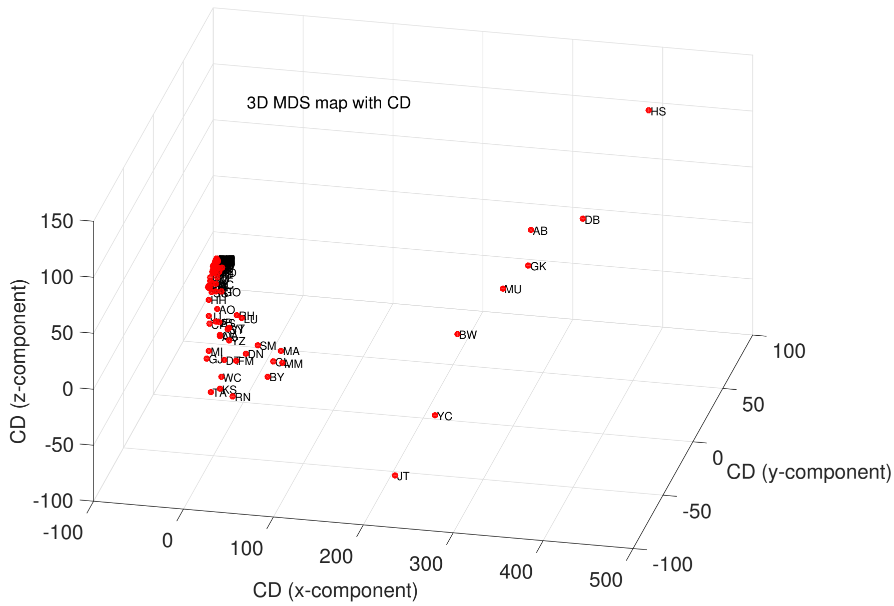

Figure 3. The three-dimensional map generated by the multidimensional scaling (MDS) with CD for 100 researchers in FC.

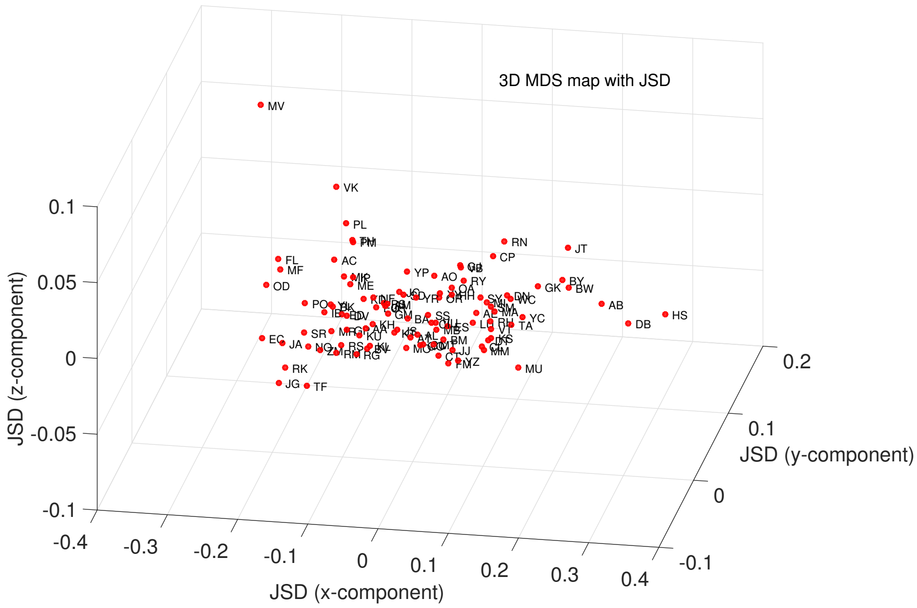

Figure 4. The three-dimensional map generated by the MDS with JSD for 100 researchers in FC.

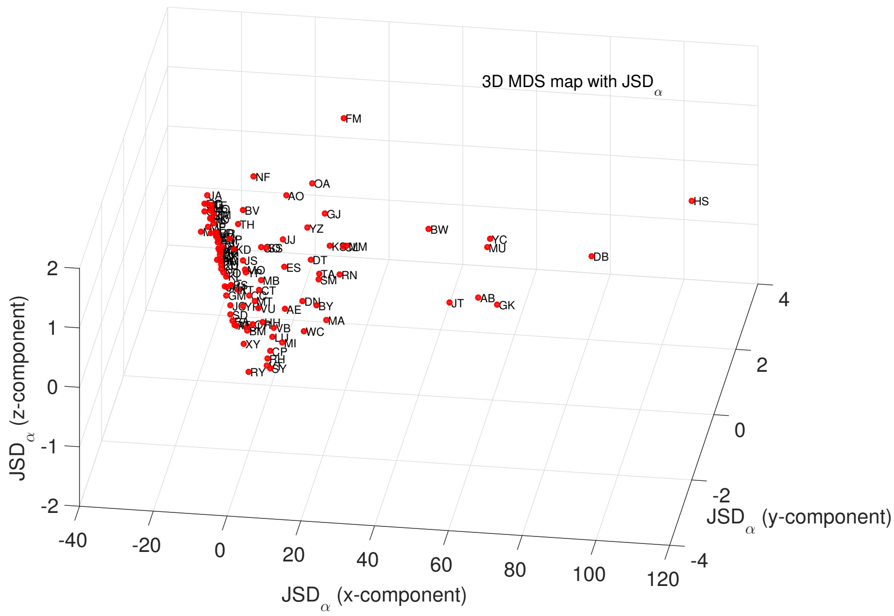

Figure 5. The three-dimensional map generated by the MDS with JSD_α, $\alpha = 0.7$, for 100 researchers in FC.

4.2. Comparing and Visualizing Scientific Output by Means of HC

For an alternative visualization of the results, we use matrix $\boldsymbol{\Delta}$ to feed an HC based on the successive (agglomerative) clustering and average-linkage method. The HC generates the hierarchical trees [52,53] shown in Figures 6–8. We can note the emergence of similar patterns for the three metrics used, which reflect the relative positioning of the FC researchers in terms of their scientific output. The fractional JSD_α has the advantage of producing a better discrimination of the patterns.

In conclusion, the MDS charts and the "trees" are alternative with different characteristics, but lead to identical conclusions. Moreover, we verify that the approach is robust in the sense that distinct metrics for quantifying dissimilarities produce charts of the same type. We also tested for profiles in terms of country, but no relevant conclusions emerged. Nonetheless, in the case of the JSD_α, supported by the tree visualization scheme, we verify the clear existence of clusters. There is no additional data available to analyze these clusters further, but an empirical educated estimation points to the effects of age and scientific subareas of research as the major issues that have influence on the results.

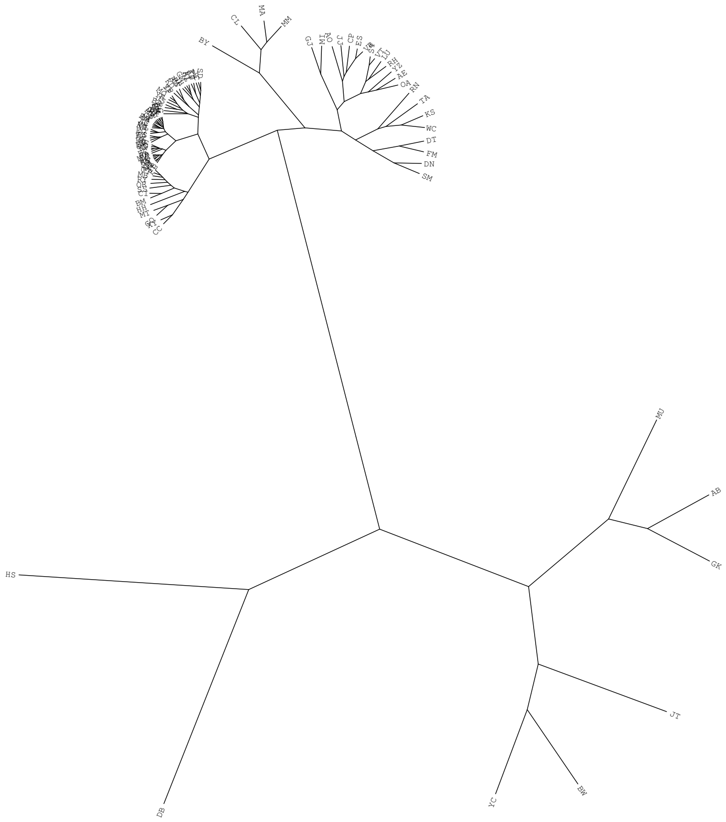

Figure 6. Hierarchical tree generated by the hierarchical clustering (HC) with CD for 100 researchers in FC.

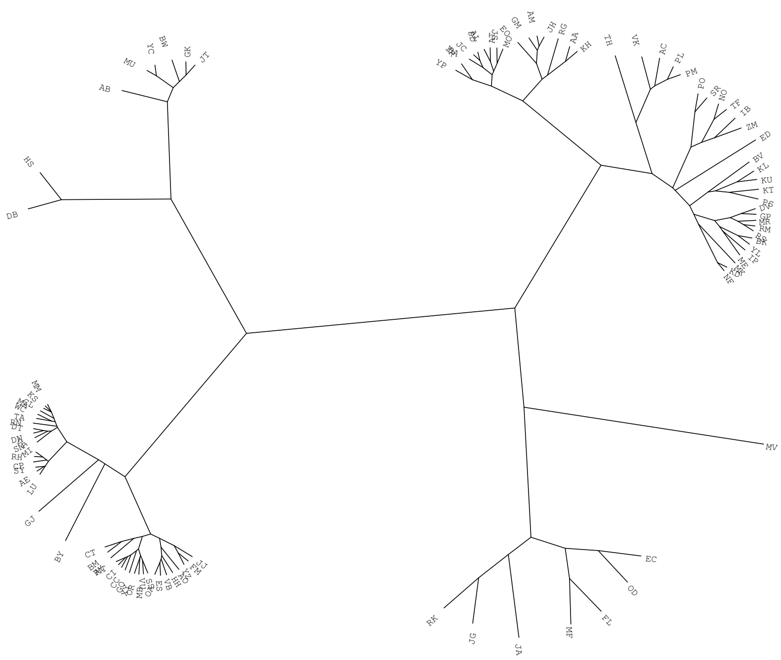

Figure 7. Hierarchical tree generated by the HC with JSD for 100 researchers in FC.

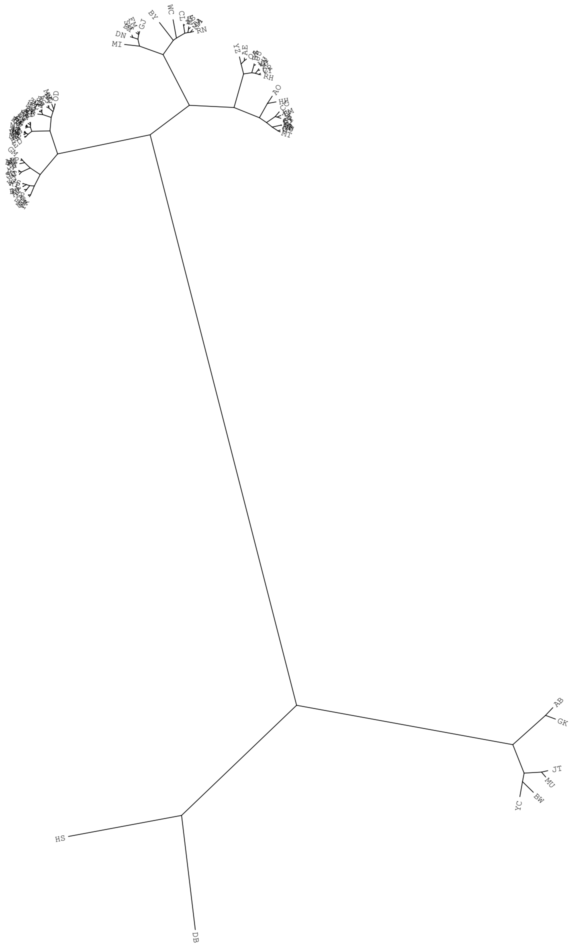

Figure 8. Hierarchical tree generated by the HC with JSD_{α}, $\alpha = 0.7$, for 100 researchers in FC.

5. Conclusions

This paper proposed an approach to compare and visualize the scientific output of researchers in FC that takes into account the complete citation profiles. We adopted different measures for quantifying dissimilarities between citation profiles, namely, the Canberra distance and the classical and fractional Jensen–Shannon divergence. The information was visualized with the MDS and HC techniques. The charts generated provided a direct interpretation of the results in terms of the relative positioning of the researchers according to their scientific output. The fractional Jensen–Shannon divergence led to a superior discrimination of the emerging patterns.

Acknowledgments: We would like to thank Thomson Web of Science (http://apps.webofknowledge.com/) for the data.

Author Contributions: J. A. Tenreiro Machado and and António M. Lopes conceived, designed and performed the experiments, analyzed the data and wrote the paper. Both the authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bornmann, L.; Daniel, H.D. Selecting scientific excellence through committee peer review—A citation analysis of publications previously published to approval or rejection of post-doctoral research fellowship applicants. *Scientometrics* **2006**, *68*, 427–440.
2. Hirsch, J.E. An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16569–16572.

3. Bornmann, L.; Daniel, H.D. What do we know about the *h* index? *J. Am. Soc. Inf. Sci. Technol.* **2007**, *58*, 1381–1385.
4. Egghe, L. How to improve the *h*-index. *Scientist* **2006**, *20*, 15–16.
5. Van Raan, A.F. Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics* **2006**, *67*, 491–502.
6. Cronin, B.; Meho, L. Using the *h*-index to rank influential information scientists. *J. Am. Soc. Inf. Sci. Technol.* **2006**, *57*, 1275–1278.
7. Glänzel, W. On the opportunities and limitations of the *H*-index. *Sci. Focus* **2006**, *1*, 10–11.
8. Ruscio, J. Taking Advantage of Citation Measures of Scholarly Impact: Hip Hip *h* Index! *Perspect. Psychol. Sci.* **2016**, *11*, 905–908.
9. Kelly, C.D.; Jennions, M.D. The *h* index and career assessment by numbers. *Trends Ecol. Evol.* **2006**, *21*, 167–170.
10. Díaz, I.; Cortey, M.; Olvera, À.; Segalés, J. Use of *h*-index and other bibliometric indicators to evaluate research productivity outcome on swine diseases. *PLoS ONE* **2016**, *11*, e0149690.
11. Asnafi, S.; Gunderson, T.; McDonald, R.J.; Kallmes, D.F. Association of *h*-index of Editorial Board Members and Impact Factor among Radiology Journals. *Acad. Radiol.* **2017**, *24*, 119–123.
12. Braun, T.; Glänzel, W.; Schubert, A. A Hirsch-type index for journals. *Scientometrics* **2006**, *69*, 169–173.
13. Lacasse, J.R.; Hodge, D.R. Ranking Disciplinary Journals with the Google Scholar *H*-index. *J. Soc. Work Educ.* **2017**, *47*, 579–596.
14. Banks, M.G. An extension of the Hirsch index: Indexing scientific topics and compounds. *Scientometrics* **2006**, *69*, 161–168.
15. Yaminfirooz, M.; Gholinia, H. Multiple *h*-index: A new scientometric indicator. *Electron. Libr.* **2015**, *33*, 547–556.
16. Sidiropoulos, A.; Katsaros, D.; Manolopoulos, Y. Generalized Hirsch *h*-index for disclosing latent facts in citation networks. *Scientometrics* **2007**, *72*, 253–280.
17. Egghe, L. Dynamic *h*-index: The Hirsch index in function of time. *J. Am. Soc. Inf. Sci. Technol.* **2007**, *58*, 452–454.
18. Bornmann, L.; Daniel, H.D. What do citation counts measure? A review of studies on citing behavior. *J. Doc.* **2008**, *64*, 45–80.
19. Iglesias, J.E.; Pecharromán, C. Scaling the *h*-index for different scientific ISI fields. *Scientometrics* **2007**, *73*, 303–320.
20. Würtz, M.; Schmidt, M. The stratified *H*-index. *Ann. Epidemiol.* **2016**, *26*, 299–300.
21. Ausloos, M. Assessing the true role of coauthors in the *h*-index measure of an author scientific impact. *Phys. A Stat. Mech. Its Appl.* **2015**, *422*, 136–142.
22. Taber, D.F. Quantifying publication impact. *Science* **2005**, *309*, 2166–2166.
23. Egghe, L. An improvement of the *h*-index: The *g*-index. *ISSI Newslett.* **2006**, *2*, 8–9.
24. Jin, B. *H*-index: An evaluation indicator proposed by scientist. *Sci. Focus* **2006**, *1*, 8–9.
25. Kosmulski, M. A new Hirsch-type index saves time and works equally well as the original *h*-index. *ISSI Newslett.* **2006**, *2*, 4–6.
26. Zhang, C.T. The *e*-index, complementing the *h*-index for excess citations. *PLoS ONE* **2009**, *4*, e5429.
27. Harzing, A.W. *The Publish or Perish Book*; Tarma Software Research: Melbourne, Australia, 2010.
28. Jin, B.; Liang, L.; Rousseau, R.; Egghe, L. The *R*-and *AR*-indices: Complementing the *h*-index. *Chin. Sci. Bull.* **2007**, *52*, 855–863.
29. Ruane, F.; Tol, R. Rational (successive) *h*-indices: An application to economics in the Republic of Ireland. *Scientometrics* **2008**, *75*, 395–405.
30. Guns, R.; Rousseau, R. Real and rational variants of the *h*-index and the *g*-index. *J. Informetr.* **2009**, *3*, 64–71.
31. Vinkler, P. The π-index: A new indicator for assessing scientific impact. *J. Inf. Sci.* **2009**, *35*, 602–612.
32. Dorogovtsev, S.N.; Mendes, J.F. Ranking scientists. *Nat. Phys.* **2015**, *11*, 882–883.
33. Kilbas, A.A.A.; Srivastava, H.M.; Trujillo, J.J. *Theory and Applications of Fractional Differential Equations*; Elsevier: Amsterdam, The Netherlands, 2006; Volume 204.
34. Gorenflo, R.; Mainardi, F. *Fractional Calculus*; Springer: New York, NY, USA, 1997.
35. Baleanu, D.; Diethelm, K.; Scalas, E.; Trujillo, J.J. *Models and Numerical Methods*; World Scientific: Singapore, 2012; Volume 3.

36. Ionescu, C.M. *The Human Respiratory System: An Analysis of the Interplay between Anatomy, Structure, Breathing and Fractal Dynamics*; Springer: New York, NY, USA, 2013.
37. Lopes, A.M.; Machado, J. Fractional order models of leaves. *J. Vib. Control* **2014**, *20*, 998–1008.
38. Machado, J.T.; Lopes, A.M. The persistence of memory. *Nonlinear Dyn.* **2015**, *79*, 63–82.
39. Machado, J.; Lopes, A.M. Analysis of natural and artificial phenomena using signal processing and fractional calculus. *Fract. Calc. Appl. Anal.* **2015**, *18*, 459–478.
40. Machado, J.; Lopes, A.; Duarte, F.; Ortigueira, M.; Rato, R. Rhapsody in fractional. *Fract. Calc. Appl. Anal.* **2014**, *17*, 1188–1214.
41. Machado, J.; Mata, M.E.; Lopes, A.M. Fractional state space analysis of economic systems. *Entropy* **2015**, *17*, 5402–5421.
42. Lance, G.N.; Williams, W.T. Computer programs for hierarchical polythetic classification ("similarity analyses"). *Comput. J.* **1966**, *9*, 60–64.
43. Lance, G.N.; Williams, W.T. Mixed-Data Classificatory Programs I—Agglomerative Systems. *Aust. Comput. J.* **1967**, *1*, 15–20.
44. Machado, J.T. Fractional Order Generalized Information. *Entropy* **2014**, *16*, 2350–2361.
45. Valério, D.; Trujillo, J.J.; Rivero, M.; Machado, J.T.; Baleanu, D. Fractional calculus: A survey of useful formulas. *Eur. Phys. J. Spec. Top.* **2013**, *222*, 1827–1846.
46. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*; Wiley: Hoboken, NJ, USA, 2012.
47. Cox, T.F.; Cox, M.A. *Multidimensional Scaling*; CRC Press: Boca Raton, FL, USA, 2000.
48. Hartigan, J.A. *Clustering Algorithms*; Wiley: Hoboken, NJ, USA, 1975.
49. Aggarwal, C.C.; Hinneburg, A.; Keim, D.A. *On the Surprising Behavior of Distance Metrics in High Dimensional Space*; Springer: New York, NY, USA, 2001.
50. Sokal, R.R.; Rohlf, F.J. The comparison of dendrograms by objective methods. *Taxon* **1962**, *11*, 33–40.
51. Lopes, A.M.; Machado, J.T. Integer and fractional-order entropy analysis of earthquake data series. *Nonlinear Dyn.* **2016**, *84*, 79–90.
52. Machado, J.A.T.; Lopes, A.M. Analysis and visualization of seismic data using mutual information. *Entropy* **2013**, *15*, 3892–3909.
53. Lopes, A.M.; Machado, J.T. Analysis of temperature time-series: Embedding dynamics into the MDS method. *Commun. Nonlinear Sci. Numer. Simul.* **2014**, *19*, 851–871.